The New TotalView

TotalView, CUDA, ReplayEngine, MemoryScape and ThreadSpotter

Chris Gottbrath
March 1st, 2011

What's New with TotalView

ROGUE WAVE SOFTWARE

TotalView Team: Application Verification and Optimization

- Highly scalable interactive GUI debugger
  - Supports basic and advanced usage
  - Used with workstations and the largest supercomputers
  - Makes developers more productive and reduces project risks
- Powerful features to enable debugging MPI parallel programs
- Compatible with wide variety of compilers across numerous operating systems

- Parallel memory analysis and error detection
  - Intuitive for both intensive and infrequent users
- Inductive user interface
- Easily integrated into the validation process

- Analyzes memory access and thread communication
- Pinpoints performance issues and provides specific guidance
- Designed for developers that aren't performance optimization experts

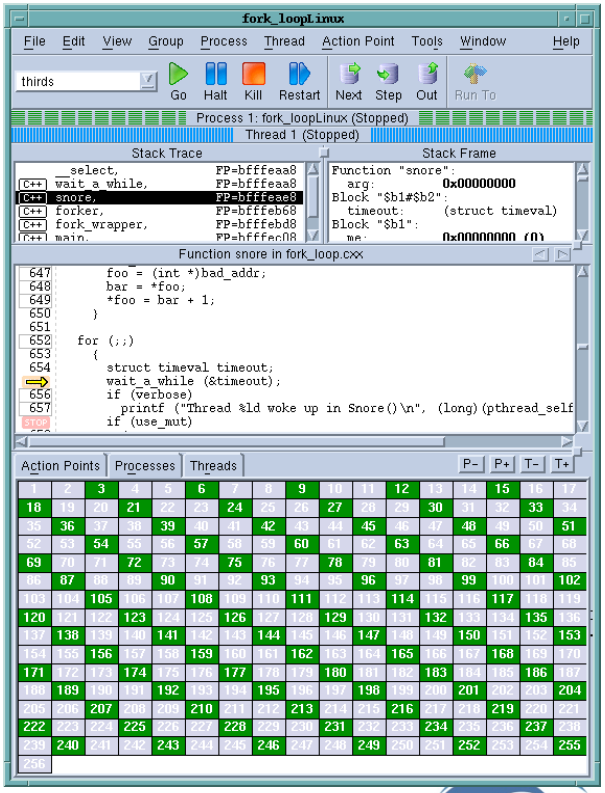
AddOn:

- Parallel record and deterministic replay
- Radically simplifies many debugging tasks
- Allows straightforward investigation of otherwise stochastic bugs

ROGUE WAVE SOFTWARE

What is TotalView?

- **Application Analysis and Debugging Tool: Code Confidently**
  - **Debug and Analyze C/C++ and Fortran on Linux, Unix or Mac OS X**
  - **Laptops to supercomputers (BG, Cray)**
  - **Makes developing, maintaining and supporting critical apps easier and less risky**



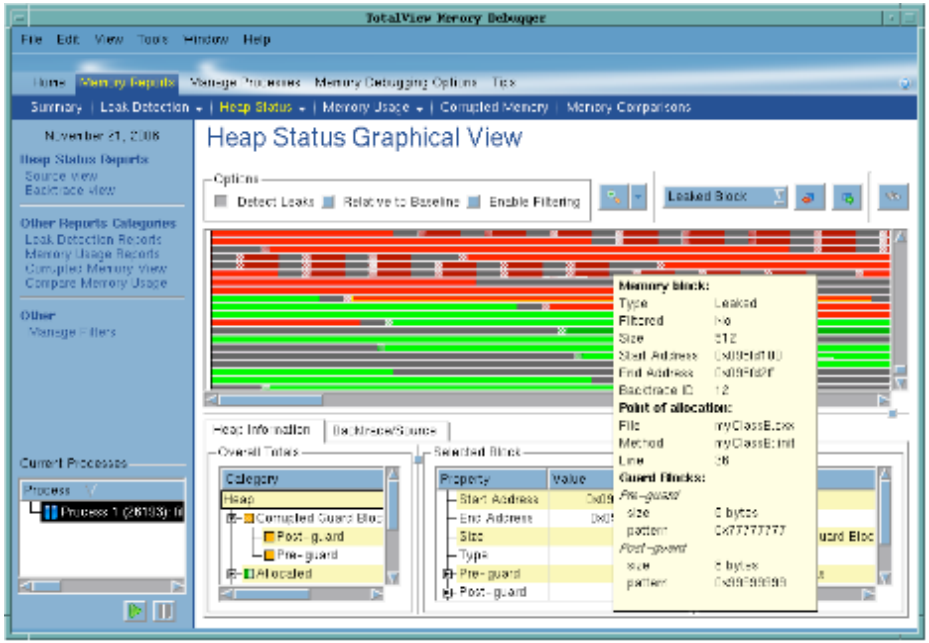
- **Major Features**
  - **Easy to learn graphical user interface with data visualization**
  - **Parallel Debugging**
    - **MPI, Pthreads, OpenMP, UPC**
    - **Optional CUDA Support**
  - **Includes a** Remote Display Client **freeing users to work from anywhere**
  - **Memory Debugging with** MemoryScape
  - **Optional Reverse Debugging with** ReplayEngine
  - **TV Team will include** ThreadSpotter **for Optimization**
  - **Batch Debugging with TVScript and the CLI**

ROGUE WAVE SOFTWARE

What Is MemoryScape?

- **Runtime Memory Analysis : Eliminate Memory Errors**
  - Detects memory leaks *before* they are a problem
  - Explore heap memory usage with powerful analytical tools
  - Use for validation as part of a quality software development process

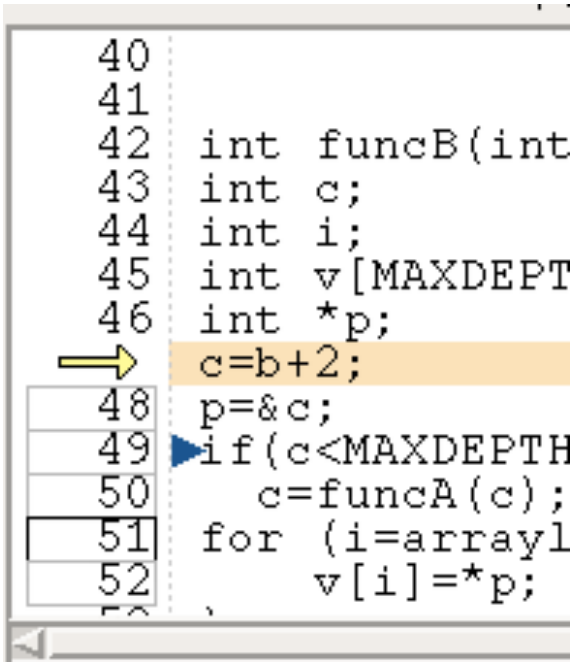
- **Major Features**
  - Detects
    - Malloc API misuse
    - Memory leaks
    - Buffer overflows
  - Supports
    - C, C++, Fortran
    - Linux, Unix, and Mac OS X
    - MPI, pthreads, OMP, and remote apps
  - Low runtime overhead
  - Easy to use
    - Works with vendor libraries
    - No recompilation or instrumentation
  - Enables Collaboration



What Is ReplayEngine?

- **Reverse Debugging Tool: Radically simplify your debugging**
  - Captures and Deterministically Replays Execution
  - Eliminate the Restart Cycle and Hard-to-Reproduce Bugs
  - Step Back and Forward by Function, Line, or Instruction

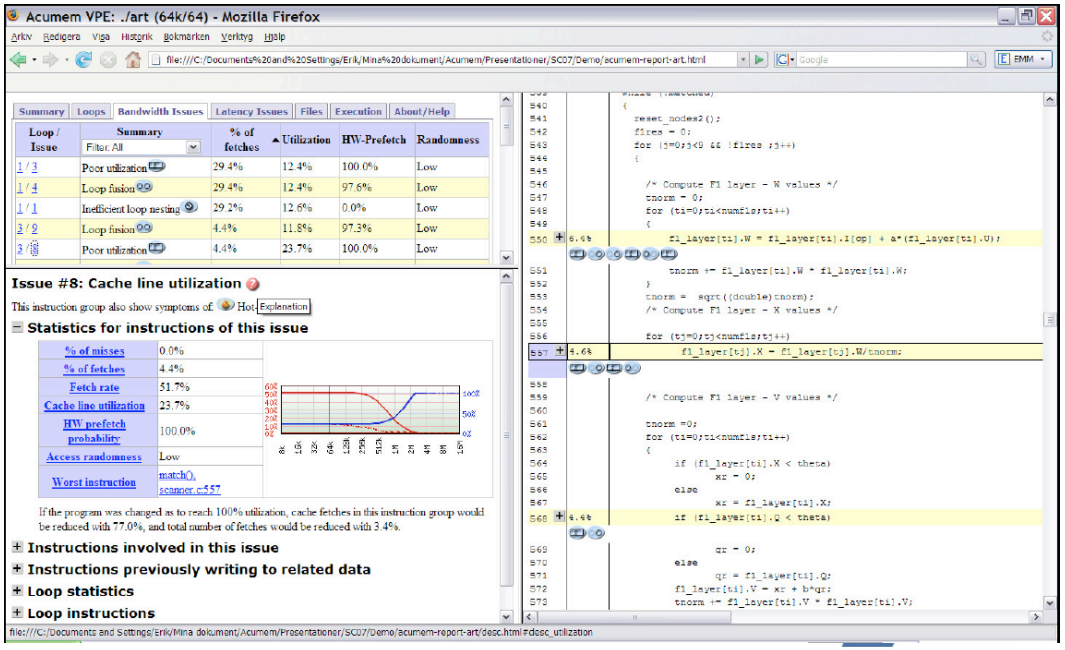
- **Major Features**
  - Simple extension to TotalView
    - No recompilation or instrumentation
    - Explore data and state in the past just like in a live process
  - Supported on Linux x86 and x86-64
  - Supports MPI, Pthreads, and OpenMP



ROGUE WAVE SOFTWARE

What is ThreadSpotter?

- **Runtime Cache Performance Optimization Tool: Tune into the Multi-Core Era**
  - **Realize More of the Performance Offered by Multi/Many-Core Chips**
  - **Quickly Detects and Prioritizes Issues -- and then Provides Usable Advice!**
    - **Brings Cache Performance Into Reach for Every Developer**
    - **Makes Experienced Cache Optimizers Hyper-Efficient**
- **Features**
  - **Supports Linux x86/x86-64**
  - **Any compiled code**
  - **Runtime Analysis**
    - **Low overhead**
  - **Cache Modeling**
    - **Prioritizes Issues**
    - **Identifies Problem Lines of Code**
  - **Provides Advice**
    - **Explanations**
    - **Examples**
    - **Detailed statistics (if desired)**



ROGUE WAVE SOFTWARE

Batch Debugging

- **Using tvscript, multiple debugging sessions can be run without the need for recompiling, unlike with printf**
- **A single compile is all that's needed, i.e.,**
  - gcc -g -o server-dbg server.c
- **tvscript syntax:**
  - tvscript [options] [filename] [-a program_args]

tvscript

- **tvscript lets you define what events to act on, and what actions to take**
- **Typical events**
  - Action_point
  - Any_memory_event
  - Guard_corruption
  - error
- **Typical actions**
  - Display_backtrace **[**-level ***level-num*****]** **[*num_levels*]** **[*options*]**
  - List_leaks
  - Save_memory
  - Print **[**-slice **{*slice_exp*]** **{*variable* | *exp*}**
- **tvscript also supports external script files, utilizing TCL within a CLI file allowing the generation of even more complex actions to events**

ROGUE WAVE SOFTWARE

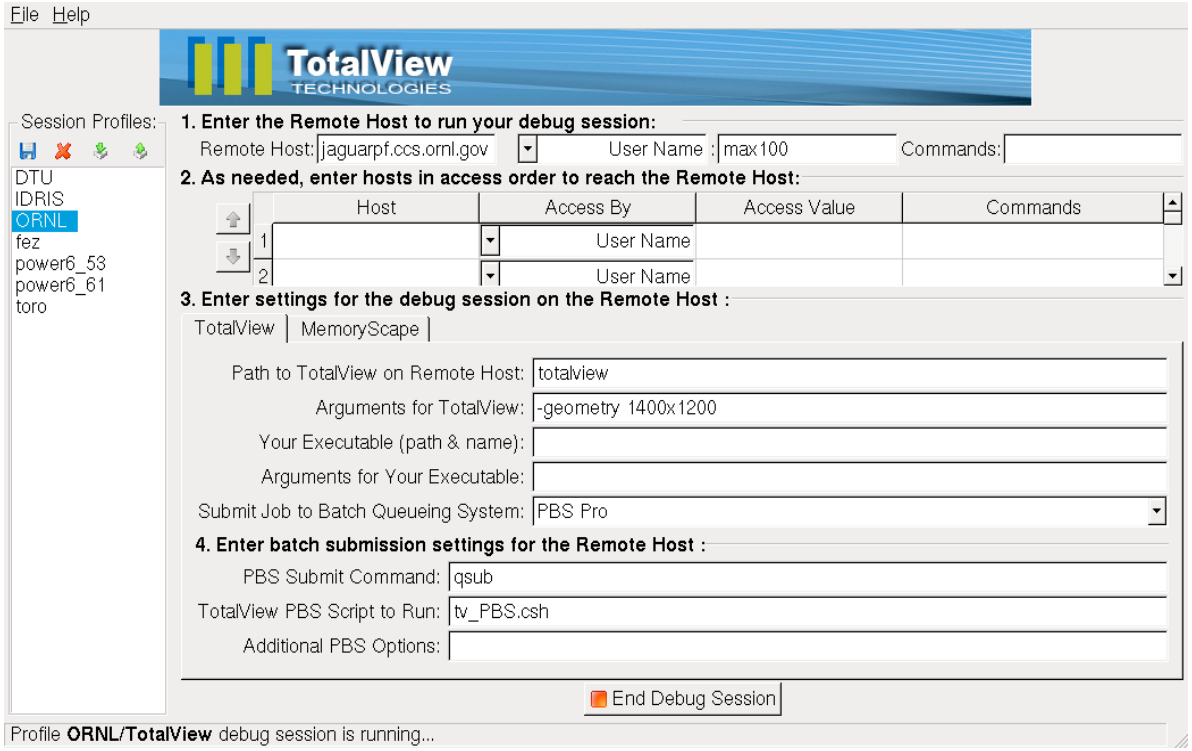
tvscript

- **Example**
- The following tells tvscript to report the contents of the *foreign_addr* structure each time the program gets to line 85
- **-create_actionpoint "#85=>print foreign_addr"**
- Typical output blocks sample with tvscript:

```
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!
! Print
!
! Process:
!    ./server (Debugger Process ID:  1, System ID:  12110)
! Thread:
!    Debugger ID:  1.1, System ID:  3083946656
! Time Stamp:
!    06-26-2008 14:04:09
! Triggered from event:
!    actionpoint
! Results:
!     foreign_addr = {
!       sin_family = 0x0002 (2)
!       sin_port = 0x1fb6 (8118)
!       sin_addr = {
!         s_addr = 0x6658a8c0 (1717086400)
!       }
!       sin_zero = ""
!     }
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!
```

ROGUE WAVE SOFTWARE

TotalView Remote Display Client



- The Remote Display Client offers users the ability to easily set up and operate a TotalView debug session that is running on another system.
- Provides for a connection that is
  - Easy
  - Fast
  - Secure
- The Remote Display Client is available for:
  - Linux x86
  - Linux x86-64
  - Windows XP
  - Windows Vista
  - Mac OS X Leopard and Snow Leopard
- The Client also provides for submission of jobs to batch queuing systems PBS Pro and Load Leveler

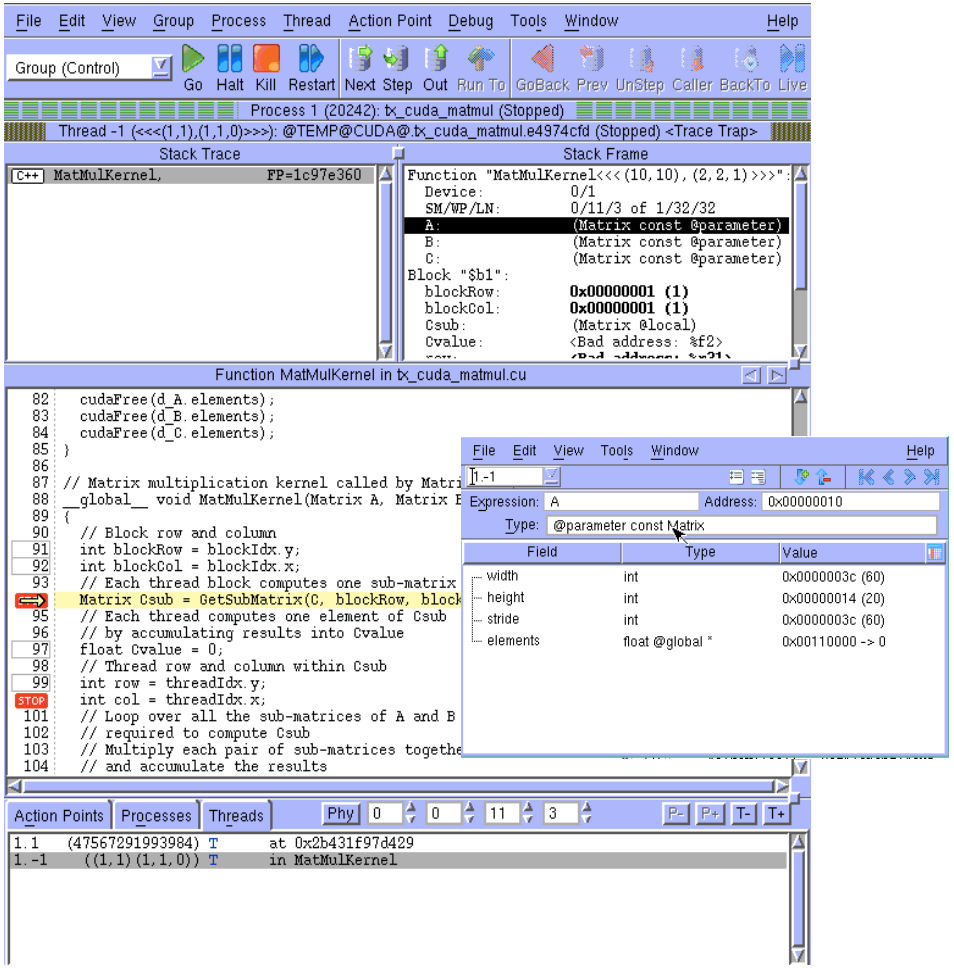
ROGUE WAVE SOFTWARE

TotalView 8.9, MemoryScape 3.1, ReplayEngine 1.8

- **What is new**
  - TotalView for CUDA (add on feature)
  - C++View
  - Multi-Dimensional Array Display
  - Parallel Backtrace
  - TVScript for BlueGene and Cray XT
  - Remote Display Client - support for Mac OS X and Windows 7
  - ReplayEngine for Infiniband (select configurations)
  - Significant Bug Fixing: 85 total, 45 user
  - Numerous Platform updates

ROGUE WAVE SOFTWARE

TotalView for CUDA

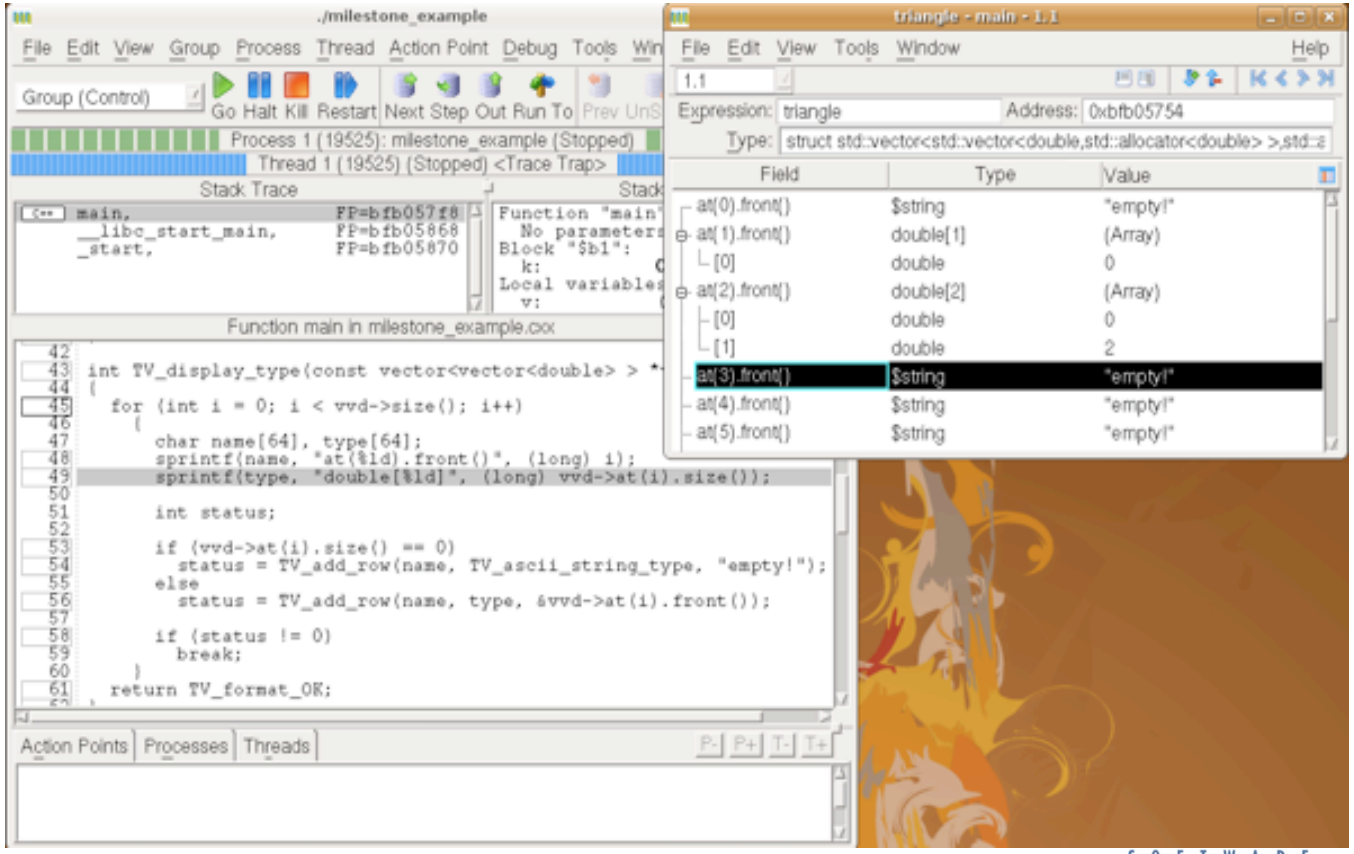


- Optional Feature (separately licensed)
- Characteristics
  - Debugging of application running on the GPU device (not in an emulator)
  - Full visibility of both Linux threads and GPU device threads
  - Fully represent the hierarchical memory
  - Thread and Block Coordinates
  - Device thread control
  - Handles CUDA function inlining
  - Reports memory access errors
  - Multi-Device Support
  - Can be used with MPI
- Supported with
  - CUDA 3.0 SDK
  - running on a Linux-x86-64 environment that is supported for both CUDA and TotalView
    - RHEL 4u8, 5u3
    - SLES 11, OpenSUSE 11.1

ROGUE WAVE SOFTWARE

C++View

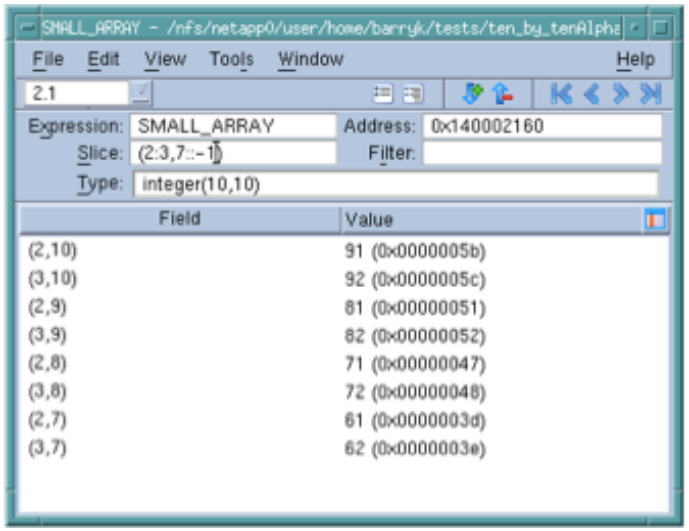
- C++View is a simple way for you to define type transformations
  - Simplify complex data
  - Aggregate and summarize
  - Check validity
- Transforms
  - Type-based
  - Compose-able
  - Automatically visible
- Code
  - C++
  - Easy to write
  - Resides in target
  - Only called by TotalView

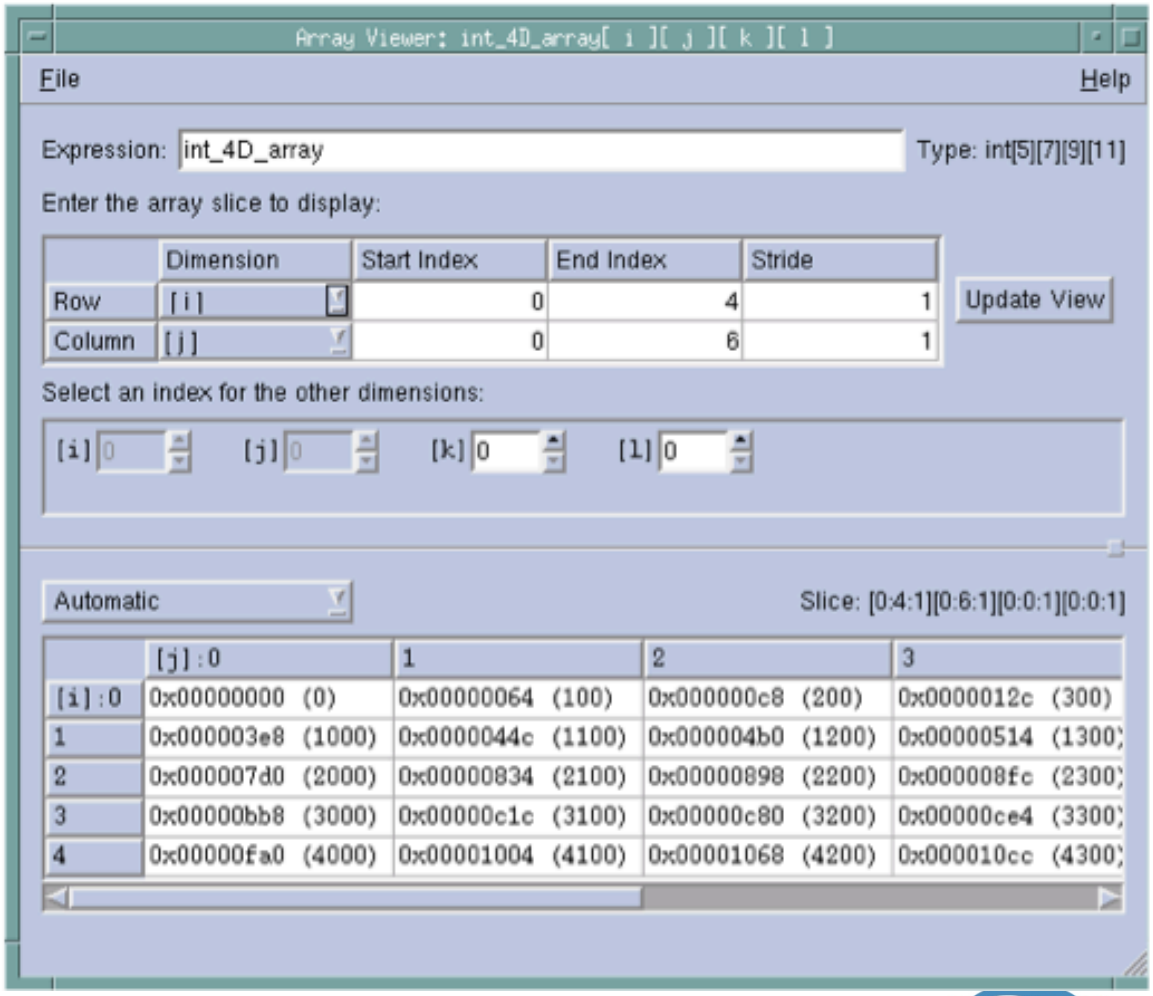


SOFTWARE

Multi-Dimensional Array Viewer

- See your arrays on a "Grid" display
- 2-D, 3-D… N-D
- Arbitrary slices
- Specify data representation
- Windowed data access
  - Fast

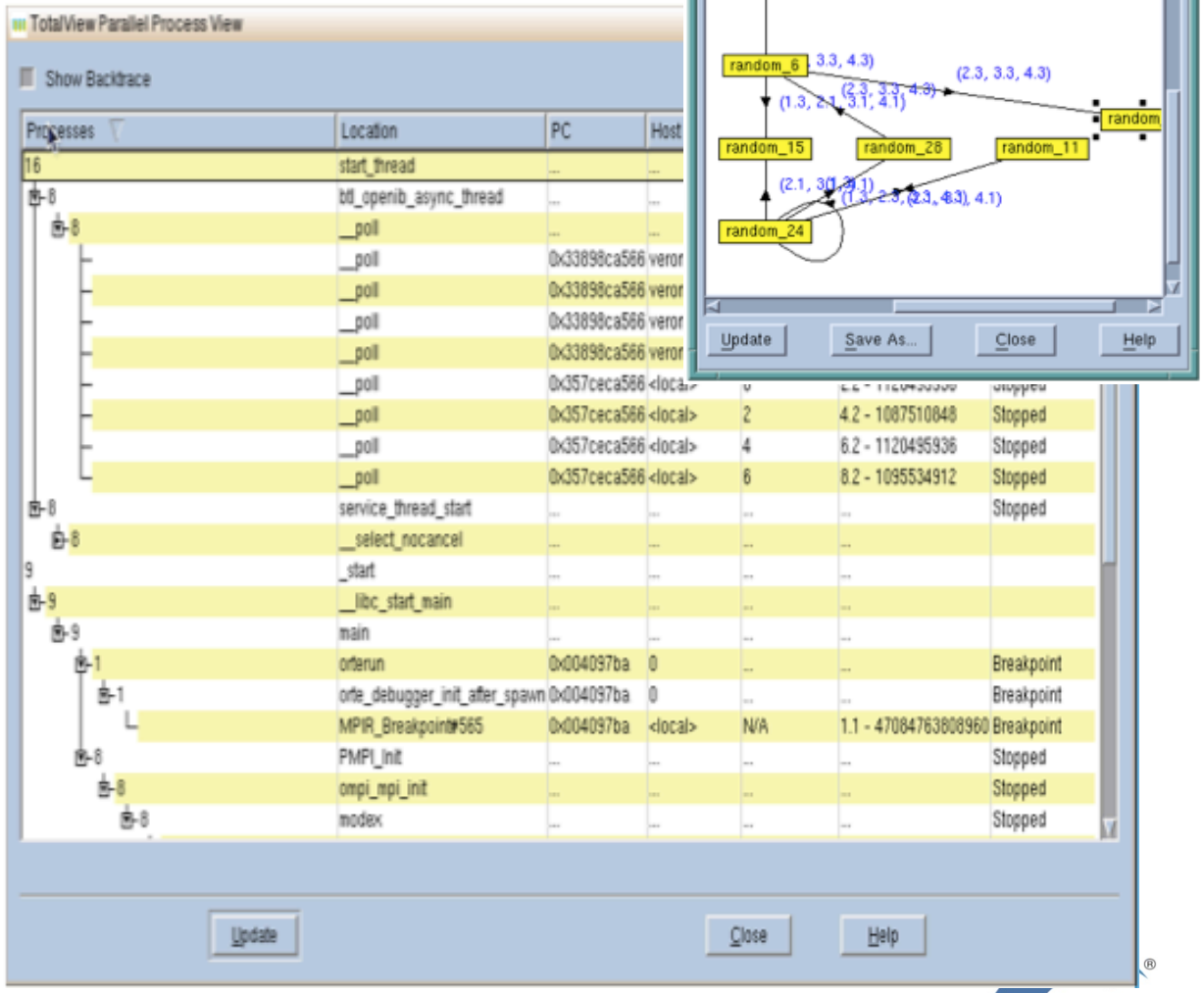




ROGUE WAVE SOFTWARE

Parallel Backtrace View

- Groups threads by common stack backtrace frames
- Starts with a compact representation of large jobs
- Text Based Tree
  - Expand/Collapse
- Elide Tree
- Shows
  - Status
  - PC
- Dive or Dive in New



ROGUE WAVE SOFTWARE

Enhanced ReplayEngine Support

- ReplayEngine
  - Provides for record and deterministic replay
  - Supported on Linux-x86 and Linux-x86-64
- Typical Cluster Interconnects
  - Gigabit Ethernet
  - Infiniband
    - Mellanox and QLogic
    - Compatibility Channel: IPoverIB
    - Channels: IBVerb, PSM (QLogic only)
    - Changing IB channels may require a recompile (MVAPICH) or a runtime switch (OpenMPI)
- ReplayEngine has extended support for Infiniband native transport channels
  - Ethernet: works with MPICH, MPICH2, Open MPI, Intel MPI, MPT
  - Mellanox Infiniband
    - IPoverIB works with Open MPI, MVAPICH, Intel MPI
    - IB verb works with Open MPI, MVAPICH
  - Qlogic Infiniband
    - IPoverIB works with Open MPI, MVAPICH, Intel MPI
    - IB verb works with Open MPI, MVAPICH
    - PSM is not supported
  - MVAPICH2 is not supported

ROGUE WAVE SOFTWARE

New Features in Next Versions

- TotalView
  - CUDA 3.2 support
    - New CUDA Registers
    - CUDA Call Stacks
    - Host-Pinned Memory Support
    - CUDA CLI Commands
  - Support for Bull MPI Environment
  - CLI Array Statistics
  - Platform Updates
    - RHEL 6, Fedora 14, GCC 4.5.2, Intel Composer XE, PGI 10.9
- ReplayEngine
  - Improved Infiniband Support
- Expected : Mid Q2

ROGUE WAVE SOFTWARE

TotalView for CUDA

ROGUE WAVE SOFTWARE

GPU Compute Accelerators

- Lots of Excitement
- Technology Trends
  - CPU Processors Multi-Core
  - GPUs have very many extremely simple cores
  - Leverages gaming/graphics market
- Multiple Vendors
  - NVIDIA Tesla and Fermi
  - AMD Firestream
- Multiple Potential Language/Runtime Choices
  - NVIDIA CUDA for C
  - OpenCL
  - PGI Accelerated Fortran
  - PGI CUDA for Fortran
  - CAPS HMPP
  - OpenMP

ROGUE WAVE SOFTWARE

NVIDIA GPU accelerator architecture

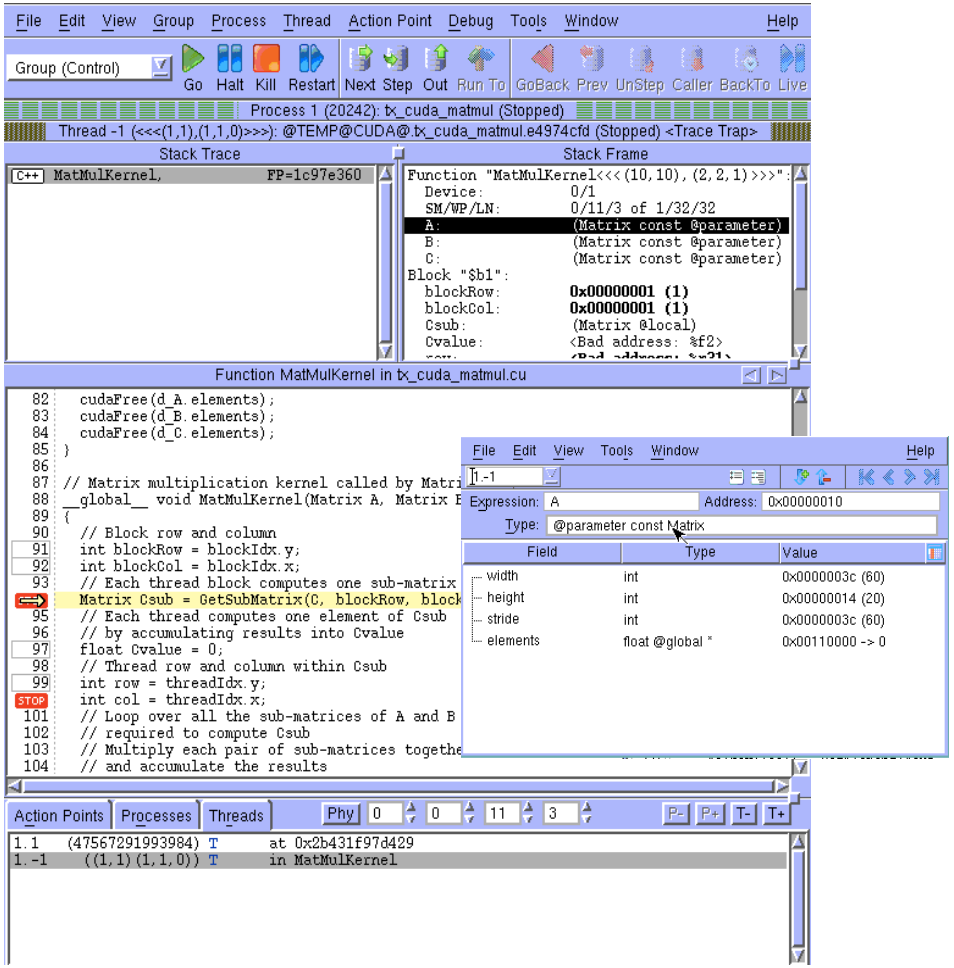
- Used in conjunction with conventional CPUs
  - Acts as an accelerator to a host process
  - Host processes may be clustered together using MPI
- Distinct processor architecture
  - Compared to host CPU
  - Features vector instructions
- Massively multi-core
  - Hundreds of streaming multiprocessors
  - Potentially 10k+ thread contexts
- Hierarchical memory with more layers
  - Local (thread)
  - Shared (block)
  - Global (GPU)
  - System (host)

Programming for the GP-GPU

- CUDA
  - Function-like kernels are written for the calculations to be performed on the GPU
    - Data parallel style, one kernel per unit of work
  - Presents a hierarchical organization for thread contexts
    - 2D grid of blocks
    - 3D block of thread
  - Exposes memory hierarchy explicitly to the user
  - Includes routines for managing device memory and data movement to and from device memory using streams
- Programming challenges
  - Coordinating CPU code + device code
  - Understanding what is going on in each kernel
    - Exceptions
  - Understanding memory usage
  - Understanding performance characteristics

ROGUE WAVE SOFTWARE

TotalView for CUDA

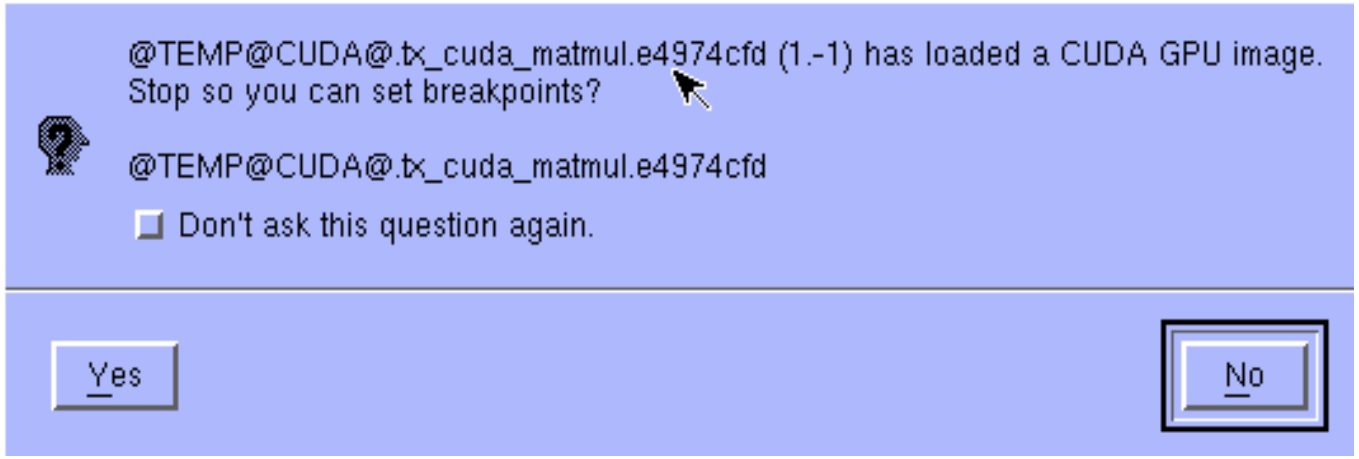


- Optional Feature (separately licensed)
- Characteristics
  - Debugging of application running on the GPU device (not in an emulator)
  - Full visibility of both Linux threads and GPU device threads
  - Fully represent the hierarchical memory
  - Thread and Block Coordinates
  - Device thread control
  - Handles CUDA function inlining
  - Reports memory access errors
  - Multi-Device Support
  - Can be used with MPI
- Supported with
  - CUDA 3.0 SDK
  - running on a Linux-x86-64 environment that is supported for both CUDA and TotalView
    - RHEL 4u8, 5u3
    - SLES 11, OpenSUSE 11.1

ROGUE WAVE SOFTWARE

Separate Image

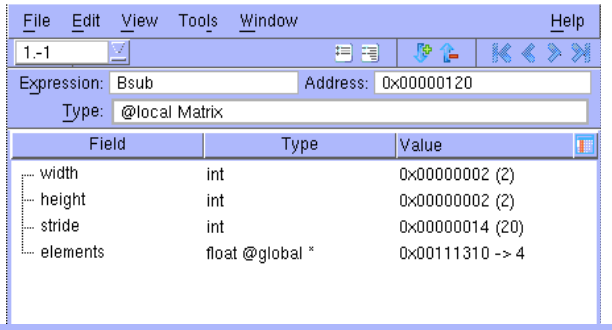
- **When a new kernel is loaded you get the option of setting breakpoints**



ROGUE WAVE SOFTWARE

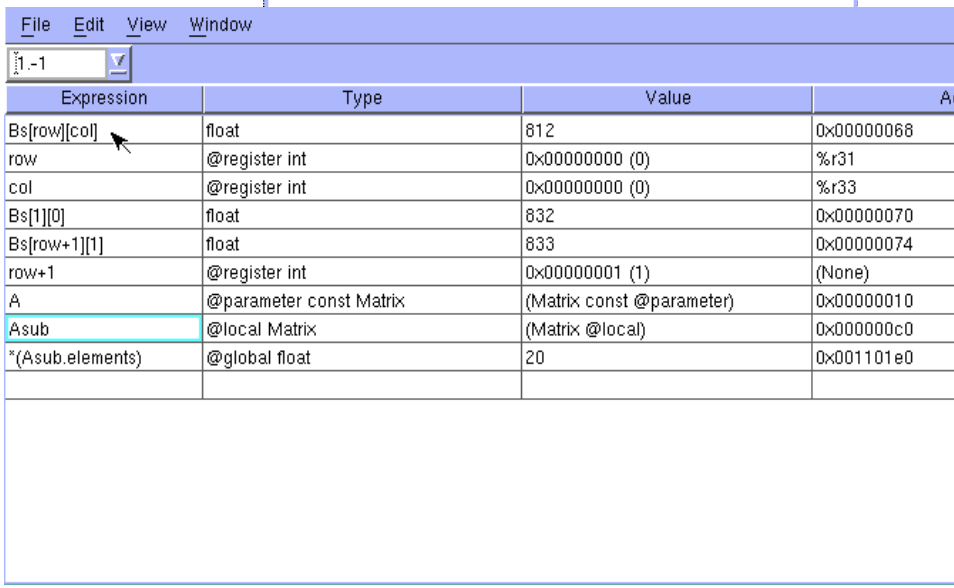
Storage Qualifiers

- **Denotes location in hierarchical memory**
  - **Part of the type**
  - **Each memory space has a separate address space so 0x00001234 could mean several places**
- **Used throughout expression system**
  - **You can cast to switch between different spaces**



Storage Qualifier	Meaning
@parameter	Address is an offset within parameter storage.
@local	Address is an offset within local storage.
@shared	Address is an offset within shared storage.
@constant	Address is an offset within constant storage.
@global	Address is an offset within global storage.
@register	Address is a PTX register name (see below).

Figure 7. Storage qualifier names



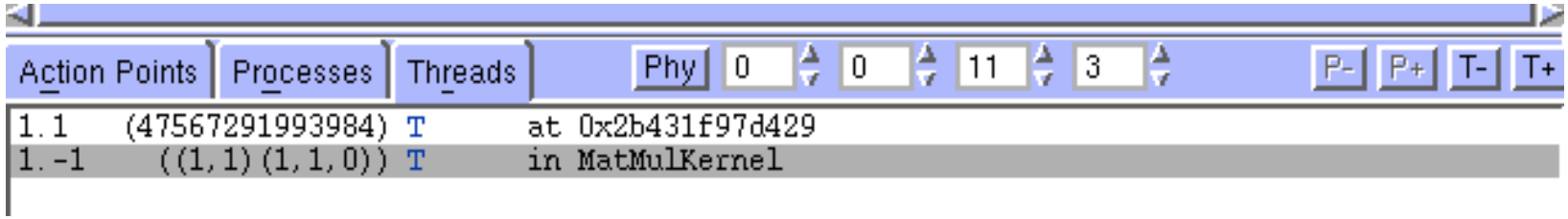
Device Threads and Warps

- **Warps advance synchronously**
  - **They share a PC**
- **Single stepping**
  - **Advances the warp containing the focus thread**
  - **Stepping over a** __syncthreads() **call advances all the relevant threads**
- **Continue and runto**
  - **Continues more than just the warp**
- **Halt**
  - **Stops all the host and device threads**

ROGUE WAVE SOFTWARE

Device Thread Navigation

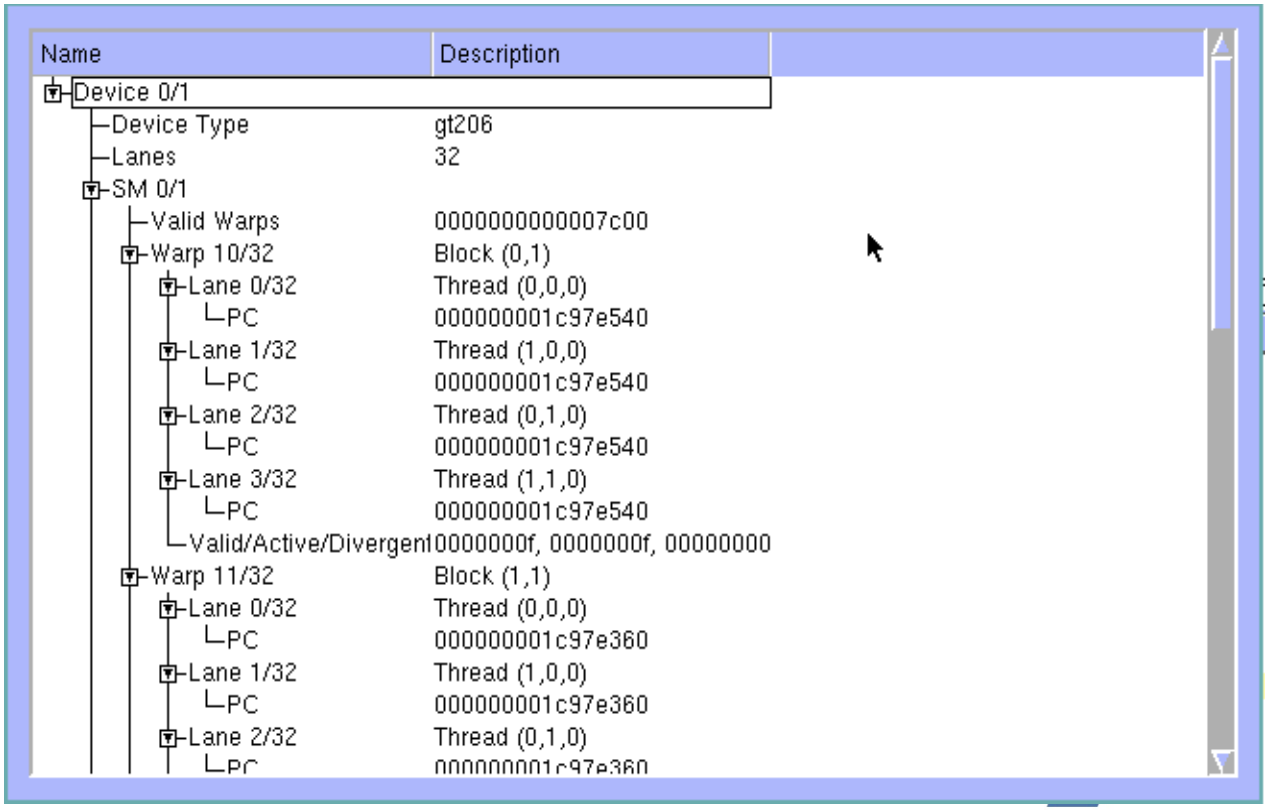
- **Two coordinate systems**
  - Hardware and Logical
    - Hardware: Device, SM, Warp, Lane
    - Logical: 2D Grid of Blocks, 3D Thread Within Grid
  - Toggle to switch input
  - Spinboxes
  - Invalid selections are refused
- **Logical coordinates are displayed elsewhere in the GUI**



ROGUE WAVE SOFTWARE

GPU Device Status Display

- Display of PCs across SMs, Warps and Lanes
- Updates as you step
- Shows what hardware is in use
- Helps you map between logical and hardware coordinates



ROGUE WAVE SOFTWARE

Debugging CUDA applications

- Applications can take advantage of
  - Kernels execute asynchronously
    - Overlap of communication and computation
    - The same kernel can operate on multiple streams
  - Multi-process applications
  - Utilization of multiple GPUs at the same time
  - Multi-level parallelism
    - MPI + OpenMP + CUDA
- The debugger should support
  - Codes that use the full capabilities of CUDA
  - Troubleshooting problems that might behave differently based on the relative timing of asynchronous events
  - This will require advanced debugging interfaces in the CUDA runtime environment
- Interface rapidly developing with each release of the SDK
  - Rogue Wave is working closely with NVIDIA to take advantage of capabilities as they are introduced

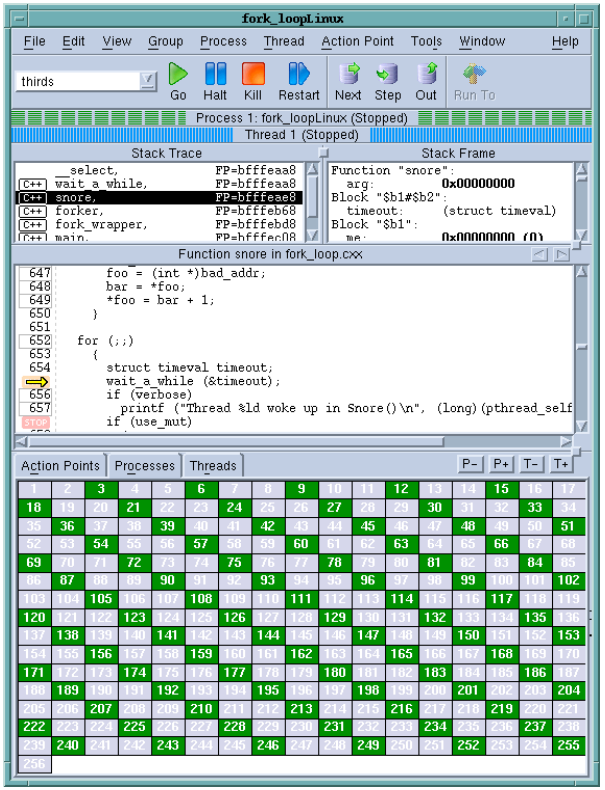
ROGUE WAVE SOFTWARE

Resource Slides

ROGUE WAVE SOFTWARE

What is TotalView?

- **What is TotalView?**
  - **Program Analysis, Debugging and Optimization Tool**
  - **For developers working with C/C++ and Fortran on Linux, Unix or Mac OS X**
  - **Workstations to Supercomputers (BG, Cray)**
  - **Makes developing, maintaining and supporting critical applications easier and less risky**
- **Major Features**
  - **Easy to learn graphical user interface with data visualization**
  - **Parallel Debugging**
    - **MPI, Pthreads, OpenMP, UPC**
    - **Optional CUDA Support**
  - **Includes a** Remote Display Client **freeing users to work from anywhere**
  - **Memory Debugging with** MemoryScape
  - **Optional Reverse Debugging with** ReplayEngine
  - **TV Team will include** ThreadSpotter **for Optimization**
  - **Batch Debugging with TVScript and the CLI**



ROGUE WAVE SOFTWARE

How can TotalView help you?

Debugging means examining a specific controlled instance of program execution

Provides an answer to the question : “What is my program *really* doing?”

- Threads and/or MPI
  - **When you have**
    - **Deadlocks and hangs**
    - **Race conditions**
  - **It provides**
    - **Asynchronous thread control**
    - **Powerful group mechanism**
- Fortran and/or C++
  - **Complex data structures**
    - **Diving and recursive dive**
  - **STL Collection Classes**
    - **STLView**
  - **Rich class hierarchies**
    - **Powerful type-casting features**
- Memory Analysis
  - **Leaks and Bounds Errors**
    - **Automatic error detection tools**
  - **Out of Memory Errors**
    - **Analysis of heap memory usage by file function and line**
- Data Analysis
  - **Numerical errors**
    - **Extensible data visualization**
    - **Slicing and filtering of arrays**
    - **Powerful expression system**
    - **Conditional watchpoints**

ROGUE WAVE SOFTWARE